Pergamon

Vision Res., Vol. 36, No. 15, pp. 2297–2302, 1996

Printed in Great Britain
0042-6989/96 $15.00 + 0.00
0042-6989(95)00259-6

Relative Efficiency for the Detection of Apparent Motion*

BRYAN L. GROS,†‡ DAVID R. POPE,‡ THEODORE E. COHN‡

Received 26 April 1995; in revised form 7 September 1995

We measured the relative efficiency for motion and position discriminations of brief, localized spot stimuli with a technique that makes no assumptions about sites of noise or information loss in the visual system. In one task, the observer had to discriminate whether an increment was located at one (left) or another (right) closely spaced spots. In the other task, the observer had to discriminate two successive brief increments of the left spot from a left spot increment followed by a right spot increment. Ideal observer theory predicts identical performance on the two tasks. Observers' thresholds, however, were significantly lower in the motion task at all intervals between flashes (ISIs) less than 60 msec in one observer and all ISIs less than 150 msec in two other observers ($P < 0.01$, *t*-test). We conclude that this apparent motion stimulus is seen more efficiently than a non-moving stimulus, and that the higher efficiency may be due to use of a motion sensitive channel in addition to independent position sensitive channels. Copyright © 1996 Published by Elsevier Science Ltd.

Apparent motion   Position discrimination   Relative efficiency   Psychophysics   Contrast threshold

INTRODUCTION

Much of the research in the area of motion perception has used stimuli of suprathreshold contrast and measured endpoints such as minimum displacement, maximum displacement, and motion coherence. Stimuli are generally gratings or random-dot patterns that span many degrees of visual angle (see e.g. Nakayama, 1985). Occasionally, while studying motion perception in the central retina, fixation marks are used which cover much of the fovea. These broad-field stimuli and experimental endpoints, however, are quite different from those used when studying static stimuli. For these studies, small spots of light have often been used; they can be localized on the retina. Endpoints are generally contrast thresholds or detectabilities. This paper reports on an approach to studying motion perception which draws on experiences with contrast perception and allows us to make a relatively assumption free estimate of efficiency.

Visual efficiency

Efficiency has been shown to be an informative way to look at the meaning of a measured contrast threshold. It quantifies the amount of available information that the observer has used, and consequently, the amount of information that the observer failed to use. One normally begins an efficiency calculation with a conceptual ideal observer for the task under study. An ideal observer gives a standard performance against which to compare real observer performance on different tasks. It is limited only by the noise or variability in the incoming signal. Geisler (1989) has provided an ideal observer framework to study the efficiency of various early stages in the human visual system.

Efficiency has been used for many years to study the performance of the visual system for luminance or contrast perception. Barlow (1962), for example, expanding on an idea originally developed by Rose (1948), measured the quantum efficiency of subject's rods for discrimination of increment thresholds and he concluded that the peak efficiency of the visual system occurs near detection threshold. Banks *et al.* (1987) used Geisler's (1989) approach to show that at the retina, the shape of the subject's contrast sensitivity function from 5 to 40 cpd is the same as that of the ideal observer's function. In other words, the observer's quantum efficiency is constant across this range of spatial frequencies.

Two groups of investigators have measured efficiency of the visual system for moving stimuli. One group, Watson *et al.* (1983), measured the quantum efficiency for detection of a variety of achromatic stimuli. They found that the stimulus seen with the highest efficiency is a grating of about 8 cpd moving with a temporal

*Portions of this paper were originally presented at the Association for Research in Vision and Ophthalmology annual meeting, Sarasota, Florida, 1993 (Gros *et al.*, 1993).

†To whom all correspondence should be addressed at 301 David Wilson Hall, Vanderbilt Vision Research Center, Nashville, TN 37420, U.S.A. [*Tel* 615 343-7656; *Fax* 615 343-8449; *Email* grosbl@ ctrvax.vanderbilt.edu].

‡Visual Detection Laboratory, Program in Vision Science, School of Optometry, University of California, Berkeley, CA 94720–2020, U.S.A.

frequency near 4 Hz. Human detection performance for grating stimuli of various spatial and temporal frequencies had been known to peak at a particular non-zero spatial and temporal frequency (Kelly, 1979). This peak predicts that a target with these attributes, i.e., a moving target, would be most easily detected, an idea verified by Watson *et al.* (1983). In another study, Watamaniuk (1993) measured the efficiency of observers for direction discrimination of random-dot motion. As the bandwidth of the distribution of possible direction vectors increased, the observer's performance decreased. However, the efficiency of the human observers increased, indicating that the ideal observer performance declines more rapidly than the visual system's performance as this bandwidth increases.

Role of assumptions in quantifying efficiency

When calculating and measuring efficiency, however, each experimenter has had to make certain assumptions about the visual system and the source(s) of information loss. These assumptions are required to allow one to calculate how much information is lost. A common assumption, for example, is that a constant percentage of quanta are either unavailable or not used by the visual system (Barlow, 1962). An assumption of this sort is needed in order to construct a mathematical model of detectability for the human observer. Efficiency is then some function of stimulus detectability for the real observer and that for the ideal observer. In the above example, that of a constant fraction (F) of quanta used by the human observer, detectability, quantified as d' (Green & Swets, 1966), is proportional to the square root of this fraction. Efficiency, that fraction of quanta used by the human observer, can then be estimated by the ratio of actual and ideal detectabilities squared (Tanner & Jones, 1960). However, a different assumption, as, for example, that of an information-losing internal noise, might lead to a different functional relation.

In order to minimize assumptions, our strategy [see Watson *et al.* (1983)] has been to design two tasks such that the predicted performance of the ideal observer is the same on each. With this approach, we avoided having to make assumptions because we chose not to quantify efficiency. Our measurements simply determine if human performance was the same or different in the two tasks. In addition, if the two tasks are identical spatially, any difference in performance must be due to neural factors in the photoreceptors or beyond. Any effects of the optics of the eye or the photoreceptor spacing on performance would be equivalent for the two tasks.

The result of this strategy is a qualitative indication of the relative efficiency of the observer on these tasks. Since our aim was to study the motion system, we chose one task which contained a motion stimulus and one task which did not. We asked the following question: is the human observer as efficient in discriminating luminance increment stimuli which contain motion as in discriminating luminance increment stimuli without motion?

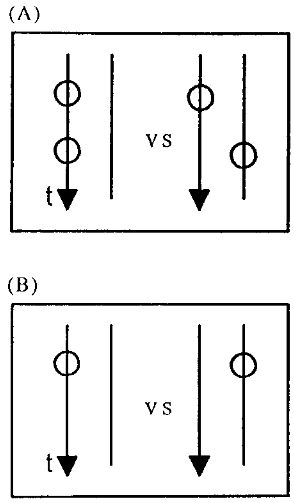

FIGURE 1. (A) Schematic of the motion discrimination task. The schematic is a space–time plot: time is on the vertical axis, increasing downward. Position in space is shown across the horizontal. The two vertical lines indicate the locations of the two spots and the open circles indicate increments of the spots in intensity. The motion discrimination task is a forced-choice discrimination between a double increment of the left spot and an apparent motion stimulus moving to the right. (B) Schematic of the position discrimination experiment. The task is a forced-choice spatial discrimination of a single increment: did the increment appear on the right or on the left?

METHODS

Tasks

Our experiment included two tasks. The first task, which we call "motion discrimination", is shown schematically in Fig. 1(A). In this figure, the vertical direction indicates time, increasing downward and the horizontal direction indicates space. An open circle indicates a luminance increment. Two spots (produced by rectangular LEDs, see below) are used, as is indicated by the vertical lines. The two stimuli in this task were either two flashes of the left spot or a single flash of the left spot followed by a flash of the right spot. The latter stimulus was the simplest motion stimulus: two-spot apparent motion. The observer's task was to indicate which stimulus was presented—the double flash, non-moving stimulus or the motion stimulus.

For an ideal observer trying to discriminate the two stimuli in the motion discrimination task, the first flash would not provide any information for the discrimination. The first flash is at the left spot in both cases. The ideal observer could safely ignore the first flash and simply determine which spot flashed during the second interval. But if the first flash is ignored, the task simply becomes a position discrimination task.

Our second task, then, is shown schematically in Fig. 1(B). We call this task "position discrimination", and the same two spots are involved. The position discrimination task uses one flash, presented either on the left or on the right. The observer's task is to indicate which spot flashed. This is the same operation that the ideal observer will perform on the first task after ignoring the first flash, so the expected performance is the same on both tasks.

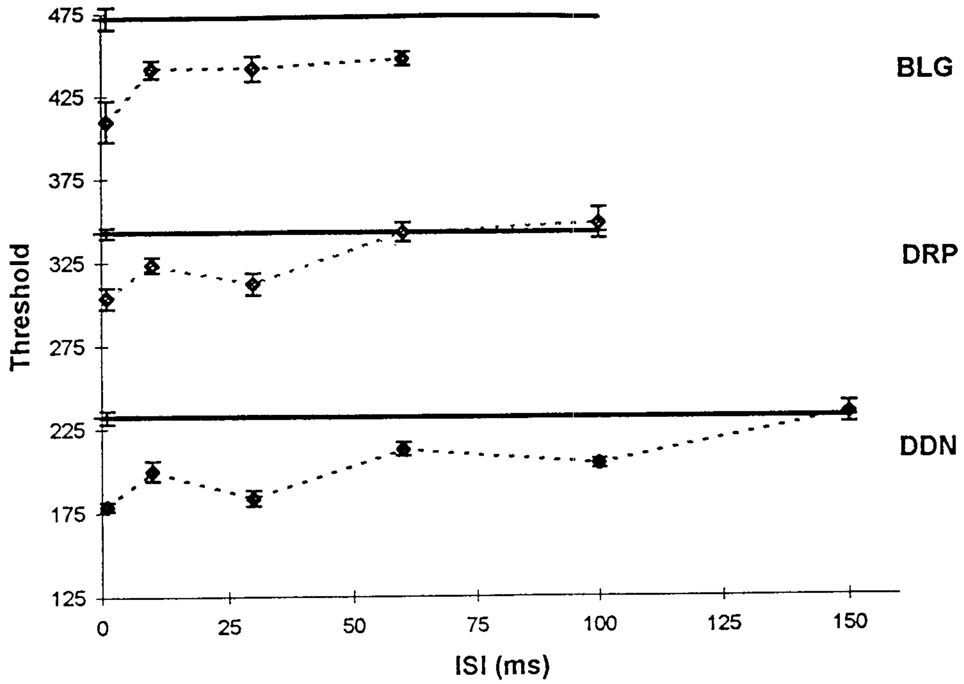

FIGURE 3. Results of the experiment for all observers. Average luminance thresholds (in arbitrary units linear with intensity) for each task are plotted against ISI in msec. The continuous horizontal line indicates the average threshold for position discrimination. Error bars represent plus and minus one standard error. Data for each subject has been displaced vertically for clarity. The thresholds for motion discrimination are significantly lower ($P < 0.01$) than the threshold for position discrimination at all ISIs tested below 60 msec for observer DRP and all ISIs tested below 150 msec for observers BLG and DDN.

DISCUSSION

The data show that subjects perform better on motion discrimination (discriminating motion from non-motion) than on position discrimination (discriminating one position from another). What explains this performance difference? We discuss several possibilities below.

Spatial or temporal uncertainty

The first flash contains no information for the discrimination of the two stimuli, and one might expect that this extra flash would only confuse the observer and hinder performance. In classical experiments, this first flash would be termed a mask. Since, in the present experiments, performance *improved* with the flash, it may be more appropriately termed a "cue" to the observer. Possibly it is a spatial cue which reduces the uncertainty as to the location of the stimulus. This idea, however, can be rejected since the spots are illuminated at a constant DC level which the observer can clearly see.

Similarly, the first flash could be a cue to the observer to reduce uncertainty as to when the stimulus will be presented. Since the subjects call for each trial with a button push, temporal uncertainty is minimal to begin with. But temporal uncertainty, however small, may be different in the two tasks due to the design.* The first flash, while containing no information for the discrimination, may cue the observer to the timing of the second, relevant flash. This idea gains support from the fall-off of thresholds as ISI increases; a longer ISI means that the potentially informative cue is further from the test flash. Consider the experiment diagrammed in Fig. 1(A). If it were changed so that, instead of the double flash being on the left it is presented on the right, the irrelevant flash would occur *after* the flash which contains the information needed for the discrimination. Any time cueing that the increment of the left spot may have added to the original motion discrimination task would be diminished, if not absent in this new configuration. If so, the thresholds for this new configuration should be larger than those measured for the original motion discrimination [Fig. 1(A)].

We tested this idea by obtaining twelve threshold estimates for each of the two motion discrimination configurations at an ISI of 10 msec. Two subjects were used. There was no significant difference in the mean thresholds for the two configurations (t-test, $P > 0.10$, both subjects). The thresholds were within 4% for one subject and 6% for the other. Therefore, we reject this hypothesis of temporal cueing.

Summation

Another related explanation for our result could be that the stimulus containing two flashes of the left spot is more visible due to summation of the two flashes over time, especially at low ISIs. This stimulus would then look brighter than the apparent motion stimulus, and the discrimination of the two, which was the essence of the task, would be easier. To evaluate whether this sort of summation affected performance, we separately mea-

*We are indebted to an anonymous referee for suggesting this possibility and the test to exclude it.

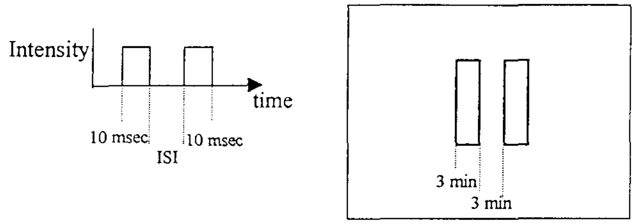

FIGURE 2. The temporal and spatial layout of the stimulus. The flashes were each 10 msec long and separated by a variable ISI. The spots were produced by LEDs, 3 arcmin wide, 9 arcmin tall, and 6 arcmin apart center-to-center.

Stimuli and apparatus

The spots were produced by rectangular light-emitting diodes (LEDs: Litronix RBG-1000) and were maintained at a visible background level during the run (approximately 25 cd/m^2). The LEDs were controlled by the current output channels of a National Instruments AT-AO-6 D/A board and the software was written in Microsoft Visual Basic and Turbo Pascal using the NIDAQ functions supplied by National Instruments. While the absolute luminance of the LEDs was not determined, a Spectra–Pritchard photometer (model UBD-10) was used to ensure that the luminance output of the LEDs was linear with the current output of the D/A board. The LEDs were mounted 57 cm from the subjects on a dark surround which was 18 deg wide by 49 deg long with a luminance of 15.45 cd/m^2 (measured with a LiteMate III photometer model 504). Beyond this surround was the darkened walls of the room which measured 2.0 cd/m^2.

The layout of the targets is shown in Fig. 2. The spots were 3 arcmin wide, 9 arcmin tall, and were separated by 6 arcmin center-to-center. The increments were 10 msec and, where applicable, separated by a variable interval (ISI). These spatial and temporal dimensions were chosen so that this would be an apparent motion stimulus and would be within the range of the reported short-range system (Anstis, 1980; Braddick, 1980). Short-range motion has been reported to activate the low-level detectors that respond to real motion, both because short-range motion produces motion after-effects (Anstis, 1980) and because it cannot be distinguished from real motion (Gregory & Harris, 1984). Six different ISIs (duration between flashes in the motion discrimination task) were tested, and their values were between 1 and 150 msec inclusive.

Procedure

The QUEST algorithm (Watson & Pelli, 1983) was used to determine the luminance threshold for each task. For each run, a task was chosen (position discrimination or motion discrimination) and an ISI value, if applicable, and two "staircases" were interleaved during each run. Typically, around twenty threshold estimates were taken for each task and each ISI value, but in all cases, there was a minimum of twelve threshold estimates. For two-alternative forced choice experiments such as the tasks described here, the threshold determined by QUEST is the 92% correct point. We assumed that fixation is constant during a trial since the subject called for each trial when he or she was ready and the duration of the stimuli were less than 200 msec.

Observers

Three observers were used, two of the authors and one undergraduate student who was naive to the purpose of the experiment. All observers had extensive practice on the tasks. Observers' heads were stabilized in a chin rest, and viewing was binocular, through natural pupils and with the observer's prescribed optical correction.

Predictions

The ideal observer prediction is that performance will be the same in both tasks. This ideal observer counts photons at the appropriate spaces and times. An alternative prediction for our tasks is that the human observer will use a system optimized for motion detection to detect the apparent motion stimuli and use a position sensitive system to detect the double flash. Current ideas on human motion perception postulate a system that is composed of receptive fields oriented in space–time and thus, optimized to detect stimuli with motion energy (Adelson & Bergen, 1985; Watson & Ahumada, 1985). Hence, performance may be different on the two tasks.

RESULTS

The results of the experiments are shown in Fig. 3. The data from all three subjects are shown in one plot, with each subject's data offset vertically for clarity. The y-axis in Fig. 3 indicates units directly proportional to absolute luminance. The x-axis indicates ISI in msec. The mean of all threshold measurements for the position discrimination task is shown by the solid, horizontal line, since there was no ISI in that task. The mean thresholds for the motion discrimination task are shown by the points connected with a dashed line. ISI is the time between flashes in that task. The error bars on each point indicate one standard error of the mean.

The data show that thresholds for motion discrimination are below those for position discrimination at all ISIs tested up to 60 msec for subject DRP and up to 150 msec for subjects BLG and DDN. A t-test on the means confirms that the threshold differences are statistically significant ($P < 0.01$) at each ISI under 60 msec for subject DRP and each ISI under 150 msec for subjects BLG and DDN.

At an ISI of 60 msec for subject DRP and 150 msec for subject DDN, the motion discrimination thresholds are virtually identical to the position discrimination thresholds. Subjects reported that there appeared to be little motion in the stimulus at these ISIs, just brightness changes at the spots. These larger ISI values place the stimulus out of the early, short-range motion category.

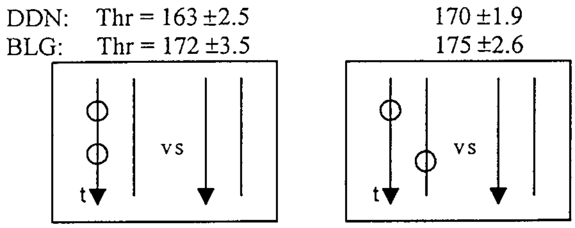

FIGURE 4. Schematic diagrams and thresholds for the detection (yes/no task) of each of the stimuli possible in the motion discrimination task. The ISI was 10 msec in each case. Thresholds and standard errors are shown above each diagram for two subjects. Schematic conventions are the same as in Fig. 1.

sured detection thresholds for both stimuli used in the motion discrimination task. The schematics for these two experiments are shown in Fig. 4. The luminance threshold was determined in each case for an ISI of 10 msec. Two subjects participated, and they reported whether the stimulus was present or absent in a yes/no task.

The data, shown in Fig. 4 above each task schematic, show that the detection thresholds for the two flash stimulus are lower than the thresholds for the motion discrimination for subject DDN. A *t*-test on the means shows this small difference is statistically significant at the 0.05 level. Subject BLG showed the opposite trend, but the difference was not significant.

The differences were very small in both cases and were not enough to account for the difference in thresholds between the motion and position discrimination tasks. These data thus reject the summation hypothesis described above, but more importantly, they allow us to draw a new inference concerning pathways in which these stimuli are processed.

Motion pathway

Since the motion and position stimuli were almost equally detectable (Fig. 4), yet were discriminable at threshold (Fig. 3), they must be detected by at least two separate, independent systems (Tanner, 1956; Thomas *et al.*, 1982). The reasoning is as follows: if a single system was transmitting signals for both stimuli at threshold, the two stimuli would not be discriminable since the higher visual areas would not know to which stimulus the system was responding.

As an example, consider the position discrimination task. The two stimuli (an increment on the left and an increment on the right) should be equally detectable if the size and duration of the increments are the same for both. Yet, as the position discrimination thresholds in Fig. 3 show, they are also discriminable from each other. At least two channels, then, must be involved in this discrimination, and their receptive fields must be separated in space, as space is the only dimension available for the discrimination. In the same way, at least three channels must be used in the motion discrimination task. In this case, there would be at least one channel sensitive to motion, in addition to the position sensitive channels involved.

Relation with ISI

As the ISI increases, thresholds for motion discrimination increase and eventually become the same as the threshold for position discrimination. This convergence in thresholds is most easily explained by the hypothesis that the motion sensitive channels play a reduced role in the discrimination and that the position sensitive channels play an increased role at longer ISIs. Consistent with this conclusion, subjects reported little impression of motion at the highest ISIs that we tested. Other studies (Braddick, 1980; Petersik, 1989) have reported that 60–100 msec is the upper temporal limit to the early or short-range motion system. Our results support this idea, and we believe that subjects are using cues other than motion to perform the discrimination when the spots are separated by longer temporal intervals.

One aspect of the data which requires comment is the exact nature of the relationship between the motion discrimination thresholds and ISI. There is no ISI value which shows a clear minimum threshold; thresholds seem to be fairly constant for ISIs less than 30 msec and then slowly rise as ISI increases to 150 msec. One might have expected the lowest thresholds for an intermediate ISI, based on data from both drifting gratings at contrast threshold (Kelly, 1979) and studies of peripheral motion perception with stimuli similar to those used in this study, but at superthreshold contrasts (Thorson *et al.*, 1969). Our failure to find a clear threshold minimum is in need of explanation. It could be due to sampling error. Alternatively, it could be due to the stimulus spatial parameters. Our stimuli are clearly localized in the fovea, whereas in the other studies, a significant portion of the stimulus, or the entire stimulus, was imaged outside of the fovea. Possibly, small areas of the fovea exhibit different temporal characteristics than the peripheral retina. Perhaps the fovea is not as selective for velocity, at least in the range of velocities we are using.

Efficiency for motion

One last question remains regarding the relative efficiency of the channel sensitive to motion. Detection performance for grating stimuli of various spatial and temporal frequencies shows a peak at a particular non-zero spatial and temporal frequency (Kelly, 1979). This peak predicts that a target with these attributes, i.e., a moving target, would be most easily detected. Watson *et al.* (1983) tested this prediction with drifting sine-wave gratings and confirmed that the human visual system detects a moving target of some moderate spatial frequency with the highest efficiency.

In our experiments, subjects showed no real difference in efficiency between detection of the two-spot apparent motion stimulus and the stationary flashed stimulus (Fig. 4). This lack of a difference in our study is surprising because of the findings of Watson *et al.* (1983) described above. We expected to see better performance on the stimulus which contained motion. Our results could be due to the difference between smooth grating motion and the sparsely sampled apparent motion. Watson *et al.*

(1983) proposed that the stimulus seen with the highest efficiency matches the weighting function (receptive field) of the most efficient detector in the visual system. If this proposal is true, we would expect the moving grating to be seen with higher efficiency than the apparent motion stimulus, since the receptive fields of the motion channel would more closely match the profile of the continuously moving grating than the sampled version of a similar stimulus. It could be that the response of the motion channels to the two flashes of the apparent motion stimulus is reduced, relative to the grating, such that it is just as sensitive as a static channel to a luminance increment.

CONCLUSION

Since Exner's experiments (cited in Sekuler *et al.*, 1990) in the nineteenth century, it has been known that a pathway exists for the detection of motion, independent of a change in luminance over time. We believe that this pathway is involved in detecting our apparent motion stimulus, even at luminance threshold. In addition, a separate pathway is involved in detecting the double flash increment stimulus, since that is discriminable from the motion stimulus at threshold. The visual system has a higher efficiency in our motion discrimination task than in a task with equal-energy stimuli that do not contain motion. We attribute this advantage to the use of motion sensitive plus motion insensitive pathways. We are currently examining conditions where the motion of spot stimuli is detected more efficiently than the non-moving stimuli.

REFERENCES

Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America, A2*, 284–299.

Anstis, S. M. (1980). The perception of apparent motion. *Philosophical Transactions of the Royal Society of London B, 290*, 153–168.

Banks, M. S., Geisler, W. S. & Bennett, P. J. (1987). The physical limits of grating visibility. *Vision Research, 27*, 1915–1924.

Barlow, H. B. (1962). A method of determining the overall quantum efficiency of visual discriminations. *Journal of Physiology, 160*, 155–168.

Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society of London B, 290*, 137–151.

Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychology Review, 96*, 267–314.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Robert E. Krieger.

Gregory, R. L. & Harris, J. P. (1984). Real and apparent movement nulled. *Nature, 307*, 729–730.

Gros, B. L., Cohn, T. E., MacLeod, D. I. A. & Pope, D. R. (1993). Relative efficiency for detection of an apparent motion stimulus. *Investigative Ophthalmology and Visual Science, 34*, 1054.

Kelly, D. H. (1979). Motion and vision II. Stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America, 69*, 1340–1349.

Nakayama, K. (1985). Biological image motion processing: a review. *Vision Research, 25*, 625–660.

Petersik, J. T. (1989). The two-process distinction in apparent motion. *Psychological Bulletin, 106*, 107–127.

Rose, A. (1948). The sensitivity performance of the human eye on an absolute scale. *Journal of the Optical Society of America, 38*, 196–208.

Sekuler, R., Anstis, S., Braddick, O. J., Brandt, T., Movshon, J. A. & Orban, G. (1990). The perception of motion. In Spillman, L. & Werner, J. S. (Eds), *Visual perception: the neurophysiological foundations* (pp. 205–230). San Diego: Academic Press.

Tanner, W. P. (1956). Theory of recognition. *Journal of the Optical Society of America, 28*, 882–888.

Tanner, W. P. & Jones, R. C. (1960). The ideal sensor system as approached through signal detection theory and the theory of signal detectability. In Morris, A. & Horne, E. P. (Eds), *Visual search problems*. Washington D.C.: National Academy of Sciences–National Research Council.

Thomas, J. P., Gille, J., & Barker, R. A. (1982). Simultaneous visual detection and identification: Theory and data. *Journal of the Optical Society of America, 72*, 1642–1651.

Thorson, J., Lange, G. D. & Biederman-Thorson, M. (1969). Objective measure of the dynamics of a visual movement illusion. *Science, 164*, 1087–1088.

Watamaniuk, S. N. J. (1993). Ideal observer for discrimination of the global direction of dynamic random-dot stimuli. *Journal of the Optical Society of America, A10*, 16–28.

Watson, A. B. & Ahumada, A. J., Jr (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A, 2, 322–341.*

Watson, A. B., Barlow, H. B. & Robson, J. G. (1983). What does the eye see best? *Nature, 302*, 419–422.

Watson, A. B. & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception and Psychophysics, 33*, 113–120.

Acknowledgements—We thank Dr Donald I. A. MacLeod for his role in the conception of these experiments and for later discussion. We also thank Dien Nguyen for his time and help, and Ben Backus and two anonymous reviewers for comments on earlier drafts of this manuscript. This paper was also submitted in partial fulfilment of the requirements of the Ph.D. degree in Physiological Optics at the University of California, Berkeley, U.S.A. This research was supported by NIH grant EY07606 to TEC, core grant EY03716 to R. Freeman, and contracts from the National Research Council (IDEA-16) and Caltrans (65W341). Publication of this document does not necessarily indicate acceptance by the National Academy of Sciences, by the Federal Transit Administration, by the Federal Highway Administration, or by the California Department of Transportation of the findings, conclusions or recommendations, either inferred or specifically expressed herein.